Perfect sampling of stationary rewards of Markov chains

Jean-Marc Vincent

Perfect sampling of stationary rewards of Markov chains

Jean-Marc Vincent

Laboratoire Informatique de Grenoble, Projet Mescal,**
51, avenue Jean Kuntzmann, 38330 Montbonnot France
(e-mail: Jean-Marc.Vincent@imag.fr)

Abstract. This article illustrates how reward backward coupling improves simulation complexity for the estimation of stationary rewards. Bounds on the coupling time for $M/M/1/C$ are given and experimental results on a large queueing network validate the practical interest of such an approach.
Keywords. Perfect sampling, coupling time, simulation efficiency.

1 Introduction

Markov chains are basic tools to study random dynamical systems on a state space $\mathcal{X}$. They play the central role of linear part of the dynamic and capture most of the dynamic characteristics. When the system is finite, the Markov chain is described by its transition kernel, (a stochastic matrix). When the system is homogeneous in time, irreducible and aperiodic, the left eigenvector π associated to the eigenvalue 1 of the transition kernel captures most of informations needed in practical applications. In real situations, the stationary distribution is analyzed via a reward function R, the interesting information for application purpose are expectation of the reward function in stationary regime $\mathbb{E}_\pi R$.

The classical methodology consists in first the computation of π by formal or numerical techniques and then compute the expected reward. Difficulties arise when the size of the system is too large so that traditional linear algebra tools could not be used.

Such systems are usually encoded by a transition function Φ and an innovation process $\{\xi_n\}_{n\in\mathbb{N}}$. The dynamic of the system with initial state x_0 is obtained by

$$\begin{aligned} X_0 &= x_0; \\ X_{n+1} &= \Phi(X_n, \xi_{n+1}),\ \text{for } n \geq 0. \end{aligned} \tag{1}$$

For a large state space, simulation provides methods based on an algorithmic representation Φ of the chain and offers new possibilities for the statistical

** This work was partially supported by ANR blanche SMS and ANR SETIN Checkbound

estimation of π. The *forward* simulation consists in an ergodic sampling on one trajectory of the chain and the mean stationary reward is estimated by

$$\hat{R}_n = \frac{1}{n} \sum_{i=0}^{n-1} R(x_n); \tag{2}$$

where $\{x_n\}_{n \in \mathbb{N}}$ is the simulated values of the stochastic recursive sequence defined by equation (1).

The main difficulty with forward simulation is to build effective confidence interval. The first reason is that the process is not in a stationary state at the beginning of the simulation. Then to avoid the dependence from the initial state we need to let the process run a sufficiently long time, the *burn in period*, and after consider the process stationary. Fixing this period is known to be a difficult problem, it depends on the spectral gap of the Markov chains, usually unknown, and bounds on the spectra of the chain are not sufficiently tight. The second reason is the correlation between the samples of the trajectory The central limit theorem could not be rigorously applied and provides only an approximation of the error. To avoid this problem replication techniques provide independent sample but with a computational overhead that could be very important.

Perfect simulation techniques have been developed in the last 10 years (see Propp (1996)). These methods guarantee the convergence to steady-state in a finite number of steps and help for the simulation control. As it will be shown in the paper, the first scheme could be improved to estimate more efficiently stationary rewards on the stationary chain. In particular, when the state space is partially ordered and the Markov chain as the reward function monotonous, the improvement is sufficient to estimate rare events probability in a "brute force" simulation.

Throughout the paper, we consider Markov chains on large state space but finite, these chains are supposed to be ergodic (irreducible and aperiodic). In the second section the perfect reward sampling is detailed and coupling time is defined. In section 3 we give the reward coupling time upper-bounded for $M/M/1/C$ queues. Experimental results illustrate the efficiency of the simulation reduction time on typical finite queuing networks.

2 Perfect simulation

The general principle of the algorithm is the following. Consider an doubly infinite sequence of independent identically distributed innovations $\{\xi_n\}_{n \in \mathbb{Z}}$ in a finite space $\mathcal{E}$. For readability purposes, the transition function $\Phi(x, \xi)$ is extended to sequences of innovations

$$\Phi(x, \xi_{i \to j}) = \Phi(\Phi(\cdots \Phi(x, \xi_i), \cdots, \xi_{j-1}), \xi_j);$$

and also extended to set of states, for $\mathcal{Z} \subset \mathcal{X}$

$$\Phi(\mathcal{Z}, \xi_{i \to j}) = \{\Phi(x, \xi_{i \to j}), \text{ such that } x \in \mathcal{Z}\}.$$

Definition 1 (Coupling). The transition function is said to be coupling if there exist a finite sequence of innovations $\{\xi_1^c, \cdots \xi_p^c\}$ with positive probability such that

$$\left|\Phi(\mathcal{X}, \xi_{1\to p}^c)\right| = 1.$$

The coupling property of the transition function is equivalent to the fact that the process $\{|\Phi(\mathcal{X}, \xi_{1\to n})|\}_{n\in\mathbb{N}}$ converges almost surely to 1. We denote by τ the coupling time associated to an innovation sequence :

$$\tau = \inf\left\{n \in \mathbb{N}; |\Phi(\mathcal{X}, \xi_{1\to n})| = 1\right\}.$$

It could be shown that τ is stochastically bounded by a geometric distribution with rate the probability of the coupling pattern. The main idea of the perfect sampling scheme is the contraction of the process Φ on sets of states.

Theorem 1 (Coupling from the past scheme). *If the transition function is coupling, the sequence of sets*

$$\Phi(\mathcal{X}, \xi_{-n\to 0})$$

converges in a finite number of steps to a singleton which is stationary distributed. The number of steps τ^* needed for convergence (backward coupling time) has the same distribution as τ.

This result obtained by Borovkov and Foss (1994) has been developed in an algorithm framework by Propp (1996) which is efficient when the transition function is monotonous. In that case, denote by $\mathcal{X}^{ext}$ the set of extremal elements $\mathcal{X}$, the simulation algorithm (1) generates one sample stationary distributed.

Algorithm 1 Backward sampling algorithm

```
k ← 1
repeat
  n ← 2^k {The simulation is started from the past with a doubling scheme}
  Z ← Φ(X^ext, ξ_{-n→0})
  {The sequence ξ_n is stored separately to be reused directly in the next loop}
until |Z| = 1
return Z
```

The doubling scheme ensures that the complexity in the number of transition function evaluation is bounded by $|\mathcal{X}^{ext}|.\mathbb{E}\tau.2$. The memory complexity (storage of the sequence ξ) is bounded by $2\mathbb{E}\tau$ and could be reduced to $\mathbb{E}\log_2 \tau + 1$. This makes the algorithm of practical interest to simulate complex monotonic systems such as interactive systems of particles, queueing networks, geometric processes, etc. When the aim is to study stationary monotonous rewards, the scheme is improved by taking the reward of the current set as in algorithm (2)

Algorithm 2 Backward reward stationary sampling algorithm

$k \leftarrow 1$
repeat
  $n \leftarrow 2^k$ {The simulation is started from the past with a doubling scheme}
  $\mathcal{Z} \leftarrow \Phi(\mathcal{X}^{ext}, \xi_{-n\to 0})$ {The sequence ξ_n is stored separately }
until $|R(\mathcal{Z})| = 1$
return $R(\mathcal{Z})$

Because the stopping condition of algorithm 2 is weaker than in algorithm 1 the associated coupling time of algorithm 2 τ_R^* verify

$$\tau_R^* \leq \tau^* \text{almost surely.}$$

It is clear that if is the reward function is strictly increasing coupling on the reward function is sufficient to ensures coupling $\tau_R^* = \tau^*$. Complexity reduction appears when the reward is concerned by a part of the global system (marginals, subsets,...). The difficulty is to estimate that complexity reduction.

3 Queueing networks

Consider now a finite capacity queueing network with K queues. Denote by C_i the capacity of queue i and λ_i^0 the external Poisson arrival rate at queue i. Consider probabilistic routing with overflow (a client that cannot enter a full queue is rejected outside), and we consider the global system as a Poisson driven process with discrete jumps at the epochs of the Poisson process with rate Λ.

In most cases, such a network is not product form and the computation algorithms are not sufficient to solve even simple systems. This comes from the rejection policy and the combinatorial explosion of the state space.

It has been shown by Glasserman and Yao (1994) that such systems are monotone. Then we could apply the perfect simulation scheme. In Vincent and Marchand (2004) and Vincent (2005), this technique was fruitfully applied to compute stationary reward distribution. To estimate the global coupling time we first study the coupling time for a simple $M/M/1/C$ queue. In Dopper et al. (2006a) the following theorem states that

Theorem 2 (Coupling time for the $M/M/1/C$). *The mean coupling time $\mathbb{E}\tau^b$ of a $M/M/1/C$ queue with arrival rate λ and service rate μ, is bounded by*

$$\begin{array}{lll} \textit{Critical bound:} & & \mathbb{E}\tau^* \leq \frac{C^2+C}{2}. \\ \textit{Heavy traffic Bound:} & \lambda > \mu, & \mathbb{E}\tau^* \leq \frac{C}{\lambda-\mu} - \frac{\mu(1-\left(\frac{\mu}{\lambda}\right)^C)}{(\lambda-\mu)^2}. \\ \textit{Light traffic bound:} & \lambda < \mu, & \mathbb{E}\tau^* \leq \frac{C}{\mu-\lambda} - \frac{\lambda(1-\left(\frac{\lambda}{\mu}\right)^C)}{(\mu-\lambda)^2}. \end{array}$$

The critical bound corresponds to the "recurrent null" case, when π is the uniform distribution. Otherwise, the distribution is concentrated around C at the top for $\lambda > \mu$ or at 0 for $\lambda < \mu$. Then the coupling time is bounded above linearly by $C\Lambda/(\mu - \lambda)$.

Consider now a monotonous reward on a simple queue based on a threshold $R_\theta(x) = \mathbb{I}_{x \geq \theta}$. Such rewards are used to estimate rare events like overflow or blocking probability.

Theorem 3 (Coupling time for reward R_θ of the $M/M/1/C$). *The reward coupling R_θ time τ_θ^b of a $M/M/1/C$ queue with arrival rate λ and service rate μ, is stochastically bounded by the the same distribution of the global coupling time of a $M/M/1/C'$ with $C' = \min\{\theta, C - \theta\}$.*

The proof is based on the 2 dimensional joint process indicating the evolution of two trajectories issued from 0 and C. Following Dopper et al. (2006*a*) the state space of the chain is described by figure 1, with $p = 1 - q = \lambda/(\lambda + \mu)$.

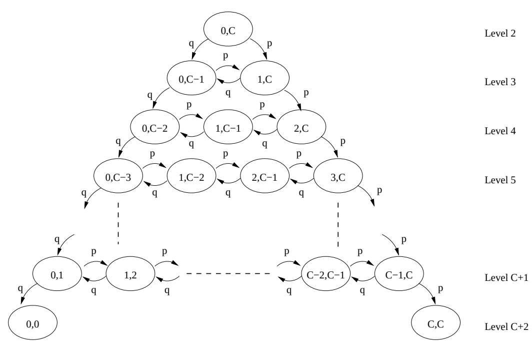


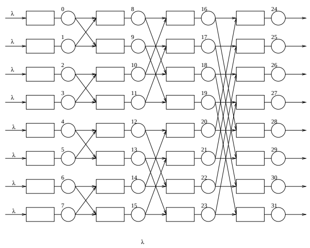
Fig. 1. Simultaneous evolution of upper and lower trajectory

The global coupling time is obtained when the process hits one of the two states $(0, 0)$ or (C, C). For the reward coupling time, the process stops at level $\theta + 1$. Then the process process is equivalent to the $M/M/\theta$. The complete computation of theorem 2 could be found in Dopper et al. (2006*b*) and theorem 3 is deduced from a coupling argument on the process which is embedded in the global process of the $M/M/1/C'$. This bound could easily be improved by considering light or heavy traffic assumptions.

4 Example

Consider a complex queueing network depicted in figure 2.

This example illustrates the efficiency of the reward coupling. In the case when the system is heavily loaded (around the recurrent null behavior) the time complexity reduction could be very important.



The estimations of the saturation probability are based on 10^6 size samples with the doubling period scheme.

We consider the saturation probability at level 4 for a homogeneous trafic with identical service rates $\mu = 1$ and a global input rate λ, and a capacity of $C = 10$ for each queue. The reward function is the saturation test on queue 31 at the last level.

λ	$\mathbb{E}\tau^b_{C-1}$	$\sigma(\tau^b_{C-1})$	$\mathbb{E}\tau^b$	$\sigma(\tau^b)$	$\mathbb{P}(sat)$
0.9	168	437	920	568	$2.3\,10^{-2}$
0.6	148	367	665	319	$2.3\,10^{-3}$
0.2	100	216	346	133	$< 10^{-6}$

Fig. 2. Delta network

5 Conclusion

Usage of reward coupling in the backward perfect simulation could improve drastically the simulation time complexity. This has been applied to queueing networks and other stochastic models like models of call centers or models of Grid schedulers. In all of these cases, even if we could not compute or bound the time reduction analytically, we observe a time complexity reduction.

For stable feed forward queueing networks, Dopper et al. (2006*a*) establish bounds on the coupling time of the network which are linear in the number of queues and at most quadratic in the queues capacities. The computation of the reward coupling time in this case remains an open problem as the problem of networks with loops.

References

Borovkov, A. and Foss, S. (1994): Two ergodicity criteria for stochastically recursive sequences. *Acta Appl. Math.*, 125–134.

Dopper, J., Gaujal, B. and Vincent, J.-M. (2006*a*): Bounds for the coupling time in queueing networks perfect simulation. *in* Numerical Solutions for Markov Chain (NSMC06), Charleston, 117–136.

Dopper, J., Gaujal, B. and Vincent, J.-M. (2006*b*): Bounds for the coupling time in queueing networks perfect simulation. Technical Report 5828, INRIA.

Glasserman, P. and Yao, D. (1994): *Monotone Structure in Discrete-Event Systems*. John Wiley and Sons.

Propp, J.and Wilson, D. (1996): Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms 9*, 223–252.

Vincent, J.-M. (2005): Perfect simulation of monotone systems for rare event probability estimation. *in* Winter Simulation Conference, Orlando.

Vincent, J.-M. and Marchand, C. (2004): On the exact simulation of functionals of stationary markov chains. *Linear Algebra and its Applications*, *386*, 285–310.